

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 15:51:26 ; Search time 11.65 seconds  
(without alignments)  
39.883 Million cell updates/sec

Title: US-09-727-198-5  
Perfect score: 57  
Sequence: 1 XPAPMKFFTTX 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt.40:\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	35	61.4	304 1 YR06_CAEEL	Q09423 caenorhabdi
2	35	61.4	327 1 DHB1_HUMAN	P14061 homo sapien
3	34	59.6	278 1 T2D7_DROME	Q27272 drosophila
4	34	59.6	453 1 TBB2_GEOCN	P32925 geotrichum
5	33	57.9	45 1 PSAT_MASLA	O31128 mastigoclad
6	33	57.9	371 1 LIG2_PHACH	P49012 phanerocha
7	33	57.9	371 1 LIG5_PHACH	P11543 phanerocha
8	33	57.9	372 1 LIG3_PHACH	P21764 phanerocha
9	33	57.9	372 1 LIG6_PHACH	P50622 phanerocha
10	33	57.9	372 1 LIG8_PHACH	P06181 phanerocha
11	33	57.9	372 1 LIGA_PHACH	P1837 phanerocha
12	33	57.9	372 1 LIGB_PHACH	P1838 phanerocha
13	33	57.9	455 1 A2AC_CAVPO	O60476 cavia porce
14	33	57.9	462 1 A2AC_HUMAN	P18823 homo sapien
15	33	57.9	476 1 PUT1_YEAST	P09368 saccharomyc
16	33	57.9	750 1 PTP2_YEAST	P29461 saccharomyc
17	33	57.9	939 1 EAE_ECO27	P19809 escherichia
18	33	57.9	969 1 NMR2_YEAST	P35724 saccharomyc
19	33	57.9	1213 1 CHS3_CANAL	P30573 candida alb
20	32	56.1	85 1 YRBA_HAETN	P45026 haemophilus
21	32	56.1	431 1 HMDH_BORBU	O51628 borrelia bu
22	32	56.1	458 1 A2AC_MOUSE	O01337 mus musculu
23	32	56.1	458 1 A2AC_MOUSE	P22086 rattus norv
24	32	56.1	499 1 VLI_HPV13	O02273 human papil
25	32	56.1	501 1 VLI_HPV13	P50820 human papil
26	32	56.1	502 1 VLI_PCPIV1	Q02274 pygmy chimp
27	32	56.1	603 1 SYD_AQUAE	O67589 aquilex aeo
28	32	56.1	730 1 PSAB_SYNY3	P29255 synechocyst
29	32	56.1	732 1 PSAB_SYNP2	P17155 synechococc
30	32	56.1	801 1 VP5_RDVA	P14583 rice dwarf
31	32	56.1	801 1 VP5_RDVF	O65437 rice dwarf
32	32	56.1	1597 1 MK4_MOUSE	O08648 mus musculu
33	32	56.1	1607 1 MK4_HUMAN	G95614 homo sapien

34	31.5	55.3	1337 1 YDM5_SCHPO	P87136 schizosacch
35	31	54.4	161 1 VG41_HSVSA	O01027 herpesvirus
36	31	54.4	295 1 YPF5_SCHPO	O14210 schizosacch
37	31	54.4	344 1 DHB1_MOUSE	P51656 mus musculu
38	31	54.4	344 1 DHB1_RAT	P51657 rattus norv
39	31	54.4	382 1 CRTG_CHLRE	O97723 chlorobium
40	31	54.4	452 1 Y397_CLOPE	P50487 clostridium
41	31	54.4	570 1 SYG_ARCFU	O29346 archaeoglob
42	31	54.4	622 1 FET5_YEAST	P43561 saccharomyc
43	31	54.4	674 1 HEL5_THEAC	O9h1x7 thermoplasm
44	31	54.4	1325 1 G160_MOUSE	P55937 mus musculu
45	31	54.4	1822 1 YM68_CAEEL	P34529 caenorhabdi

## ALIGNMENTS

```

RESULT 1
ID YR06_CAEEL STANDARD: PRT; 304 AA.
AC Q09423:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 34.8 kDa protein R0763.6 in chromosome II.
GN R0763.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL NZ;
RA Connell M.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U23452; AAK31546.1; -
DR WormPep: R0763.6; CE02025.
KW Hypothetical protein.
FT DOMAIN 92
FT DOMAIN 132
FT DOMAIN 140
FT DOMAIN 211
FT SEQUENCE 304 AA; 34846 MW; EAD74EB1BB988FDS CMC64;
Query Match 61.4%; Score 35; DB 1; Length 304;
Best Local Similarity 75.0%; Pred. No. 6.7;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 4 PMKFFTT 11
Db 235 PLMKFFTT 242
RESULT 2
ID DHB1_HUMAN STANDARD: PRT; 327 AA.
AC P14061:
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE Escradial 17 beta-dehydrogenase 1 (EC 1.1.1.62) (17-beta-HSD 1)
DE (Placental 17-beta-hydroxysteroid dehydrogenase) (20 alpha-
DE hydroxysteroid dehydrogenase) (20-alpha-HSD) (E2DH).

```

GN HSD17B1 OR EDH17B1 OR EDHB17 OR E17KSR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RA MEDLINE=89384667; PubMed=2779584;  
 RA Liu-The V., Labrie C., Zhao H.F., Couet J., Lachance Y., Simard J.,  
 RA Leblanc G., Labrie F.;  
 RT "Structure of two in tandem human 17 beta-hydroxysteroid  
 RT dehydrogenase genes.";  
 RL Mol. Endocrinol. 4:268-275(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=90231340; PubMed=2330005;  
 RA Liu-The V., Labrie C., Simard J., Lachance Y., Zhao H.F., Couet J.,  
 RA Leblanc G., Labrie F.;  
 RT "Structure of two in tandem human 17 beta-hydroxysteroid  
 RT dehydrogenase genes.";  
 RL Mol. Endocrinol. 4:268-275(1990).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=90328695; PubMed=2197970;  
 RA Liu-The V., Labrie C., Zhao H.F., Couet J., Lachance Y., Simard J.,  
 RA Cote J., Leblanc G., Lagace L., Berube D., Gagne R., Labrie F.;  
 RT "Purification, cloning, complementary DNA structure, and predicted  
 RT amino acid sequence of human estradiol 17 beta-dehydrogenase.";  
 RL Ann. N.Y. Acad. Sci. 595:40-52(1990).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RA MEDLINE=89031223; PubMed=2846351;  
 RA Peloteketo H., Isomaa V., Maenetausta O., Viikko R.;  
 RT "Complete amino acid sequence of human placental 17 beta-  
 RT hydroxysteroid dehydrogenase deduced from cDNA.";  
 RL FEBS Lett. 239:73-77(1988).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=93011163; PubMed=1327779;  
 RA Peloteketo H.E., Isomaa V., Viikko R.;  
 RT "Genomic organization and DNA sequences of human 17 beta-  
 RT hydroxysteroid dehydrogenase genes and flanking regions. Localization  
 RT of multiple Alu sequences and putative cis-acting elements.";  
 RL Eur. J. Biochem. 209:459-466(1992).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC Shen Y., Gibbs R.A.;  
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE OF 1-23.  
 RA MEDLINE=72238162; PubMed=5045524;  
 RA Burns D.J.W., Engel L.L., Bethune J.L.;  
 RT "Amino acid composition and subunit structure. Human placental 17 -  
 RT estradiol dehydrogenase.";  
 RL Biochemistry 11:2699-2703(1972).  
 RN [8]  
 RP SEQUENCE OF 51-67.  
 RA MEDLINE=73221216; PubMed=4719204;  
 RA Nicholas J.C., Harris J.I.;  
 RT "Human placental 17 -oestradiol dehydrogenase. Sequence of a tryptic  
 RT peptide containing an essential cysteine.";  
 RL FEBS Lett. 29:173-176(1973).  
 RN [9]  
 RP SEQUENCE OF 204-223.  
 RA MEDLINE=86159759; PubMed=3955020;  
 RA Murdock G.L., Chin C.C., Warren J.C.;  
 RT "Elucidation of the chemical nature of the steady-state intermediates  
 RT in the mechanism of carboxypeptidase A.";  
 RL Biochemistry 25:646-651(1986).  
 RN [10]

RP SEQUENCE OF 219-223.  
 RX MEDLINE=84008135; PubMed=6578212;  
 RA Murdock G.L., Chin C.C., Offord R.E., Bradshaw R.A., Warren J.C.;  
 RT "Human placental estradiol 17 beta-dehydrogenase. Identification of a  
 RT single histidine residue affinity-labeled by both 3-  
 RT bromoacetoxyestron and 12 beta-bromoacetoxy-4-estrene-3,17-dione.";  
 RL J. Biol. Chem. 258:11460-11464(1983).  
 RN [11]  
 RP SIMILARITY TO SHORT CHAIN DEHYDROGENASES.  
 RX MEDLINE=89330506; PubMed=2547159;  
 RA Baker M.E.;  
 RT "Human placental 17 beta-hydroxysteroid dehydrogenase is homologous  
 RT to Nodg protein of Rhizobium meliloti.";  
 RL Mol. Endocrinol. 3:881-884(1989).  
 RN [12]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 AND 1.7 ANGSTROMS).  
 RX MEDLINE=95393021; PubMed=7663947;  
 RA Chosh D., Pletnev V.Z., Zhu D.W., Mawrzak Z., Duax W.L., Pangborn W.,  
 RA Labrie F., Lin S.-X.;  
 RT "Structure of human estrogenic 17 beta-hydroxysteroid dehydrogenase  
 RT at 2.20-A resolution.";  
 RL Structure 3:503-513(1995).  
 RN [13]  
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).  
 RX MEDLINE=96398680; PubMed=8805577;  
 RA Breton R., Housset D., Mazza C., Fontecilla-Camps J.-C.;  
 RT "The structure of a complex of human 17beta-hydroxysteroid  
 RT dehydrogenase with estradiol and NADP+ identifies two principal  
 RT targets for the design of inhibitors.";  
 RL Structure 4:905-915(1996).  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
 RX MEDLINE=98192605; PubMed=9525918;  
 RA Mazza C., Breton R., Housset D., Fontecilla-Camps J.-C.;  
 RT "Unusual charge stabilization of NADP+ in 17beta-hydroxysteroid  
 RT dehydrogenase.";  
 RL J. Biol. Chem. 273:8145-8152(1998).  
 RN [15]  
 RP VARIANTS VAL-237 AND GLY-312.  
 RX MEDLINE=93278403; PubMed=8389226;  
 RA Normand T., Narod S., Labrie F., Simard J.;  
 RT "Detection of polymorphisms in the estradiol 17 beta-hydroxysteroid  
 RT dehydrogenase II gene at the EDH17B2 locus on 17q11-q21.";  
 RL Hum. Mol. Genet. 2:479-483(1993).  
 CC -1- FUNCTION: FAVORS THE REDUCTION OF ESTROGENS AND ANDROGENS.  
 CC ALSO HAS 20-ALPHA-HSD ACTIVITY. USES PREFERENTIALLY NADH.  
 CC -1- CATALYTIC ACTIVITY: Estradiol-17-beta + NAD(P)(+) = estrone +  
 CC NAD(P)H.  
 CC -1- PATHWAY: BIOSYNTHESIS OF ESTROGENS.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M27138; AAB16941.1; -  
 DR EMBL: X13440; CAA31792.1; -  
 DR EMBL: M36263; AAA35600.1; -  
 DR EMBL: M84472; AAB16942.1; -  
 DR EMBL: U34879; AAD05019.1; -  
 DR PIR: A36081; DEH057.  
 DR PDB: 1FDS; 12-FEB-97.  
 DR PDB: 1FDU; 27-MAY-98.  
 DR PDB: 1FDV; 27-MAY-98.  
 DR PDB: 1FDW; 27-MAY-98.



```

RESULT 4
TBW2_GEOCN ID TBW2_GEOCN STANDARD; PRT; 453 AA.
AC P32925;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tubulin beta-2 chain.
OS Geotrichum candidum (Oospora lactic).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Galactomyces.
OX NCBI_TaxID=27317;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92079883; PubMed=1836049;
RA Gold S.E., Casale W.L., Keen N.T.;
RT "Characterization of two beta-tubulin genes from Geotrichum
candidum.";
RL Mol. Gen. Genet. 230:104-112(1991).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; S69627; AAB2057.1; ALT_SEQ.
CC PIR; S18597; S18597.
CC DR InterPro; IPR002453; Beta.tubulin.
CC DR InterPro; IPR000217; Tubulin.
CC DR InterPro; IPR003008; Tubulin_FtsZ.
CC DR Pfam; PF00091; tubulin; 1.
CC DR PRINTS; PFO161; TUBULIN.
CC DR PROSITE; PS00227; TUBULIN; 1.
CC DR PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
CC DR PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
CC DR Microtubules; GTP-binding; Multigene family.
CC FT NP_BIND 142 148 GTP (POTENTIAL).
CC SQ SEQUENCE 453 AA; 50399 MW; 2EA908A0246E0371 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 453;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11
I I : I I I I
Db 261 PFPRLHFFTT 270

RESULT 5
PSAU_MASLA ID PSAU_MASLA STANDARD; PRT; 45 AA.
AC O31128;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit IX.
DE PSAU.
OS Mastigocladus laminosus (Fischereia sp.).
OC Bacteria; Cyanobacteria; Stigonematales; Fischerella.
OX NCBI_TaxID=1191;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 7605;

```

```

CC -1- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN. CATALYSES THE
CC C(ALPHA)-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.
CC -1- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.
CC -1- DEVELOPMENTAL STAGE: LIGNINASES ARE EXPRESSED DURING SECONDARY
CC METABOLISM, AND ARE TRIGGERED BY NUTRIENT LIMITATION.
CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. LIGNINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74229; AAA33735.1; -.
DR PDB; 1LLP; 08-MAR-96.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00141; peroxidase; 1.
DR PRINTS; PR00458; PEROXIDASE.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
DR PROSITE; PS00436; PEROXIDASE_2; 1.
DR Oxidoreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KM Lignin degradation; Signal; Zymogen; 3D-structure.
FT SIGNAL 1 21
FT PROPEP 22 28
FT CHAIN 29 371
FT ACT_SITE 71 71
FT ACT_SITE 75 75
FT ACT_SITE 204 204
FT DISULFID 31 43
FT DISULFID 42 313
FT DISULFID 62 148
FT DISULFID 277 345
FT CARBOHYD 285
SQ SEQUENCE 371 AA; 39329 MW; F6AAB123EDC92123 CRC64;

Query Match 57.9%; Score 33; DB 1; Length 371;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 APMMKFFT 10
Db 151 APQNMFFT 158

RESULT 7
LIG5_PHACH STANDARD; PRT; 371 AA.
ID LIG5_PHACH
AC P1543;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ligninase Ig5 precursor (EC 1.11.1.-) (Lignin peroxidase).
GN GLG5 OR LIP6.
OS Phanerochaete chrysosporium.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticiaceae; Phanerochaete.
OX NCBI_TaxID=5306;

RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=88152506; PubMed=3440521;
RA de Boer H.A., Zhang Y.Z., Collins C., Reddy C.A.;
RT "Analysis of nucleotide sequences of two ligninase cDNAs from a
RT white-rot filamentous fungus, Phanerochaete chrysosporium.";
RL Gene 60:93-102(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 24725 / BKM 1767;
RX MEDLINE=91187681; PubMed=2011531;
Gaskell J., Dieperink E., Cullen D.;

```

```

RT "Genomic organization of lignin peroxidase genes of Phanerochaete
RT chrysosporium.";
RL Nucleic Acids Res. 19:599-603(1991).
RL [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 24725 / BKM 1767;
RX MEDLINE=91153647; PubMed=1999283;
RA Zhang Y.Z., Reddy C.A., Raoooly A.;
RT "Cloning of several lignin peroxidase (LIP)-encoding genes: sequence
RT analysis of the LIP6 gene from the white-rot basidiomycete,
RT Phanerochaete chrysosporium.";
RL Gene 97:191-198(1991).
CC -1- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN. CATALYSES THE
CC C(ALPHA)-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.
CC -1- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.
CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. LIGNINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M18794; AAA33734.1; -.
DR EMBL; X55343; CA839033.1; -.
DR EMBL; M63496; AAA33739.1; -.
DR PIR; JN0117; OPTG5.
DR HSSP; P06181; 1B85.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00141; peroxidase; 1.
DR PRINTS; PR00458; PEROXIDASE.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
DR PROSITE; PS00436; PEROXIDASE_2; 1.
DR Oxidoreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KM Lignin degradation; Signal; Zymogen.
FT SIGNAL 1 21
FT PROPEP 22 27
FT CHAIN 28 371
FT ACT_SITE 70 70
FT ACT_SITE 74 74
FT ACT_SITE 202 202
FT DISULFID 30 42
FT DISULFID 41 311
FT DISULFID 61 146
FT DISULFID 275 344
FT CARBOHYD 283
SQ SEQUENCE 371 AA; 39417 MW; FOAE30C69D7EF5C2 CRC64;

Query Match 57.9%; Score 33; DB 1; Length 371;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 APMMKFFT 10
Db 149 APQNMFFT 156

RESULT 8
LIG3_PHACH STANDARD; PRT; 372 AA.
ID LIG3_PHACH
AC P21764;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ligninase Ig3 precursor (EC 1.11.1.-) (Lignin peroxidase).
GN GLG3 OR LIP.
OS Phanerochaete chrysosporium.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticiaceae; Phanerochaete.
OX NCBI_TaxID=5306;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 24725 / BKM 1767;
RA MEDLINE=91088334; PubMed=2129560;
RT Naidu P.S., Reddy C.A.;
RT "Nucleotide sequence of a new lignin peroxidase gene GL63 from the
RT white-rot fungus, Phanerochaete chrysosporium.";
RL Nucleic Acids Res. 18:7173-7173(1990).
CC -I- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN. CATALYSES THE
CC C(ALPHA)-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.
CC -I- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.
CC -I- DEVELOPMENTAL STAGE: LIGNINASES ARE EXPRESSED DURING SECONDARY
CC METABOLISM, AND ARE TRIGGERED BY NUTRIENT LIMITATION.
CC -I- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. LIGNINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51590; CAA35939.1; -.
DR PIR: S13723; OPJG3P.
DR HSSP: P06181; 1B80.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PF00141; peroxidase; 1.
DR PRINTS: PR00458; PEROXIDASE.
DR PROSITE: PS00435; PEROXIDASE_1; 1.
DR PROSITE: PS00436; PEROXIDASE_2; 1.
DR Oxidoreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KW Lignin degradation; Signal; Zymogen.
FT SIGNAL 1 21
FT PROPEP 22 28
FT CHAIN 29 372
FT ACT_SITE 71 71
FT ACT_SITE 75 75
FT ACT_SITE 204 204
FT DISULFID 31 43
FT DISULFID 62 148
FT DISULFID 277 345
FT CAROXYD 285 285
SQ SEQUENCE 372 AA; 39536 MW; DA0293178F4E2A4A CRC64;

Query Match 57.9%; Score 33; DB 1; Length 372;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 AP0MKFFT 10
Db 151 AP0MKFFT 158

RESULT 9
LIG6_PHACH STANDARD; PRT; 372 AA.
AC P50622;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ligninase L66 precursor (EC 1.11.1.-) (Lignin peroxidase).
GN GL66.
OS Phanerochaete chrysosporium.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticiaceae; Phanerochaete.
OX NCBI_TaxID=5306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 24725 / BKM 1767;
RA MEDLINE=91097609; PubMed=2268360;
RT Naidu P.S., Zhang Y.Z., Reddy C.A.;

```

```

RT "Characterization of a new lignin peroxidase gene (GL66) from
RT Phanerochaete chrysosporium.";
RL Biochem. Biophys. Res. Commun. 173:994-1000(1990).
CC -I- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN. CATALYSES THE
CC C(ALPHA)-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.
CC -I- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.
CC -I- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. LIGNINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M80213; AAA3737.1; -.
DR HSSP: P06181; 1B80.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PF00141; peroxidase; 1.
DR PRINTS: PR00458; PEROXIDASE.
DR PROSITE: PS00435; PEROXIDASE_1; 1.
DR PROSITE: PS00436; PEROXIDASE_2; 1.
DR Oxidoreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KW Lignin degradation; Signal; Zymogen.
FT SIGNAL 1 21
FT PROPEP 22 28
FT CHAIN 29 372
FT ACT_SITE 71 71
FT ACT_SITE 75 75
FT ACT_SITE 204 204
FT DISULFID 31 43
FT DISULFID 62 148
FT DISULFID 277 345
FT CAROXYD 285 285
SQ SEQUENCE 372 AA; 39259 MW; B192A93351F8CC16 CRC64;

Query Match 57.9%; Score 33; DB 1; Length 372;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 AP0MKFFT 10
Db 151 AP0MKFFT 158

RESULT 10
LIG8_PHACH STANDARD; PRT; 372 AA.
AC P06181;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ligninase H8 precursor (EC 1.11.1.-) (Lignin peroxidase).
GN LPOA.
OS Phanerochaete chrysosporium.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticiaceae; Phanerochaete.
OX NCBI_TaxID=5306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 24725 / BKM 1767;
RA MEDLINE=87173020; PubMed=3561490;
RT Tien M., Tu C.-P.D.;
RT "Cloning and sequencing of a cDNA for a ligninase from Phanerochaete
RT chrysosporium.";
RL Nature 326:520-523(1987).
RN [2]
RP REVISIONS.
RT Tien M., Tu C.-P.D.;
RL Nature 328:742-742(1987).
RN [3]

```



```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X54257; CAA3817.1; -.
DR PIR: S13564; OPUGBP.
DR HSSP: P06181; 1882.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PR00458; PEROXIDASE_1.
DR PROSITE: PS00435; PEROXIDASE_2; 1.
DR Oxioreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KW Lignin degradation; Signal; Zymogen.
FT SIGNAL 1 21
FT PROPE 22 28
FT CHAIN 29 372
FT ACT_SITE 71 71
FT ACT_SITE 75 75
FT ACT_SITE 204 204
FT DISULFID 31 43
FT DISULFID 42 313
FT DISULFID 62 148
FT DISULFID 277 345
FT CAROHND 285 285
FT CONFLICT 135 135
FT CONFLICT 161 161
FT CONFLICT 298 298
SQ SEQUENCE 372 AA; 39393 MW; 0ACE4598D6D6F09B CRC64;

Query Match 57.9%; Score 33; DB 1; Length 372;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 APMKFFET 10
Db 151 APOMNFFT 158

RESULT 12
LIGB PHACH STANDARD; PRT; 372 AA.
AC P31838;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ligninase B precursor (EC 1.11.1.1) (Lignin peroxidase).
GN LIPB.
OS Phanerochaete chrysosporium.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticiaceae; Phanerochaete.
OX NCBI_TaxID=5306;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 24725 / BKM 1767;
RX MEDLINE=91187681; PubMed=2011531;
RA Gaskell J., Diaperink E., Culen D.;
RT "Genomic organization of lignin peroxidase genes of Phanerochaete
RT chrysosporium";
RL Nucleic Acids Res. 19:599-603(1991).
CC -1- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN. CATALYSES THE
CC C(ALPHA)-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.
CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. LIGNINASE SUBFAMILY.
CC -1- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: X54257; CAA3817.1; -.
DR PIR: S13564; OPUGBP.
DR HSSP: P06181; 1882.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PR00458; PEROXIDASE_1.
DR PROSITE: PS00435; PEROXIDASE_2; 1.
DR PROSITE: PS00436; PEROXIDASE_2; 1.
KW Oxioreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KW Lignin degradation; Signal; Zymogen.
FT SIGNAL 1 21
FT PROPE 22 28
FT CHAIN 29 372
FT ACT_SITE 71 71
FT ACT_SITE 75 75
FT ACT_SITE 204 204
FT DISULFID 31 43
FT DISULFID 42 313
FT DISULFID 62 148
FT DISULFID 277 345
FT CAROHND 285 285
SQ SEQUENCE 372 AA; 39530 MW; E0DD2CF40425D9D5 CRC64;

Query Match 57.9%; Score 33; DB 1; Length 372;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 APMKFFET 10
Db 151 APOMNFFT 158

RESULT 13
A2AC CAVPO STANDARD; PRT; 455 AA.
AC Q60476;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Alpha-2C adrenoceptor receptor (Alpha-2C adrenoceptor).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathii; Cavidae.
OX NCBI_TaxID=10141;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=HARTLEY;
RX MEDLINE=96152573; PubMed=8573196;
RA Svensson S.P., Bailey T.J., Porter A.C., Richman J.G., Regan J.W.;
RT "Heterologous expression of the cloned guinea pig alpha 2A, alpha 2B,
RT and alpha 2C adrenoceptor subtypes. Radioligand binding and
RT functional coupling to a CAMP-responsive reporter gene.";
RL Biochem. Pharmacol. 51:291-300(1996).
CC -1- FUNCTION: ALPHA-2 ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC INDUCED INHIBITION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC PROTEINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U25724; AAA67076.1; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.

```









**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 15:50:51 ; Search time 33.9 Seconds  
(without alignments)  
61.237 Million cell updates/sec

Title: US-09-727-198-5  
Perfect score: 57  
Sequence: 1 XPAPMKFFPTX 12

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	66.7	120	3	013904
2	37	64.9	598	17	0973K4
3	36	63.2	225	5	017989
4	36	63.2	300	10	09SUS5
5	36	63.2	353	5	09XRP5
6	36	63.2	519	10	09AXL7
7	35	61.4	121	10	09SL50
8	35	61.4	163	4	09H7F5
9	35	61.4	166	11	09C018
10	35	61.4	189	17	09V243
11	35	61.4	191	17	057989
12	35	61.4	222	3	008278
13	35	61.4	292	6	09GME2
14	35	61.4	328	10	09SAD4
15	35	61.4	388	5	09XK73
16	35	61.4	443	16	050980

17	35	61.4	476	16	09H2G1	09hzq1 pseudomonas
18	35	61.4	810	3	012172	012172 saccharomyc
19	35	61.4	2016	5	09YL57	09yl57 drosophila
20	35	61.4	6315	2	09ADL6	09adl6 polyangium
21	34	59.6	162	10	09SKN0	09skn0 arabidopsis
22	34	59.6	253	5	09SRJ4	09srj4 drosophila
23	34	59.6	274	5	09XVQ9	09xvq9 caenorhabdi
24	34	59.6	284	10	09SIU3	09siu3 arabidopsis
25	34	59.6	336	10	048824	048824 arabidopsis
26	34	59.6	378	2	057286	057286 veillonella
27	34	59.6	486	10	09FM41	09fm41 arabidopsis
28	34	59.6	487	16	067873	067873 aquilex aeo
29	34	59.6	931	10	09M9A7	09m9a7 arabidopsis
30	33	57.9	202	4	09BX48	09bx48 homo sapien
31	33	57.9	246	5	09S5G2	09s5g2 drosophila
32	33	57.9	284	10	09LJD3	09lj3 arabidopsis
33	33	57.9	312	10	09FV16	09fv16 arabidopsis
34	33	57.9	325	16	066707	066707 aquilex aeo
35	33	57.9	362	3	001785	001785 phanerocha
36	33	57.9	371	3	001787	001787 phanerocha
37	33	57.9	371	3	006326	006326 phanerocha
38	33	57.9	455	3	094738	094738 rhizomucor
39	33	57.9	462	4	09HB49	09hb49 homo sapien
40	33	57.9	475	5	09VB80	09vb80 drosophila
41	33	57.9	491	5	09VB81	09vb81 drosophila
42	33	57.9	640	5	09VF65	09vf65 drosophila
43	33	57.9	1331	4	096PE1	096pe1 homo sapien
44	33	57.9	1967	5	09VPV1	09vpv1 drosophila
45	32.5	57.0	133	12	084462	084462 paramecium

## ALIGNMENTS

RESULT 1  
ID 013904 PRELIMINARY; PRT; 120 AA.  
AC 013904;  
DT 01-JAN-1999 (TREMBLrel. 09, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE HYPOTHELICAL 13.9 KDA PROTEIN C22A12.13 IN CHROMOSOME I.  
GN SPAC22A12.13.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=972;  
RA Pearson D., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
DR EMBL; Z99295; CAB16583.2; -;  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 10 30 POTENTIAL.  
FT TRANSMEM 48 68 POTENTIAL.  
SQ SEQUENCE 120 AA; 13903 MW; EB9431495A6004B2 CRC64;

Query Match 66.7%; Score 38; DB 3; Length 120;  
Best Local Similarity 75.0%; Pred. No. 2.5;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PAPMKFF 9  
Db 32 PAPVLFK 39

RESULT 2  
ID 0973K4 PRELIMINARY; PRT; 598 AA.

AC Q973K4;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE PUTATIVE THERMOPHILIN.  
 GN ST0894.  
 OS Sulfolobus tokodaii.  
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
 NCBI\_TaxID=111955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JCM 10545 / 7;  
 RX PubMed=11572479;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
 RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,  
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
 RA Oshima T., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic thermophilic  
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";  
 RL DNA Res. 8:123-140(2001).  
 DR EMBL: AP000984; BAB5908.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 598 AA; 68102 MW; 20E848698BB931F6 CRC64;

Query Match 64.9%; Score 37; DB 17; Length 598;  
 Best Local Similarity 75.0%; Pred. No. 20;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 PAMKFFTT 11  
 11:||||  
 DB 271 PVMRFFTT 278

RESULT 3  
 ID 017989 PRELIMINARY; PRT; 225 AA.  
 AC 017989;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE R08H2.6 PROTEIN.  
 GN R08H2.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Matthews L.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL: Z81575; CAB04632.1; -  
 DR InterPro: IPR003003; 7TM\_chemorecept\_2.  
 DR InterPro: IPR000168; 7TM\_nematode.  
 DR Pfam: PF01604; 7tm\_5; 1.  
 SQ SEQUENCE 225 AA; 25500 MW; BFF5FE0D5313D4 CRC64;

Query Match 63.2%; Score 36; DB 5; Length 225;  
 Best Local Similarity 60.0%; Pred. No. 12;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 PAMKFFTT 11  
 11:||||

DB 77 PCPAEFFTT 86  
 RESULT 4  
 ID 09S055 PRELIMINARY; PRT; 300 AA.  
 AC 09S055;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)  
 DE HYPOTHEICAL 34.8 KDA PROTEIN.  
 GN T17P15.210.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Choisme N., Robert C., Brothier P., Wincker P., Cattolico L.,  
 RA Artiguenave F., Saurin W., Weissenbach J., Mewes H.W., Mayer K.F.X.,  
 RA Lemcke K., Schueller C., Quetier F., Salanoubat M.;  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL049658; CAB41147.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 300 AA; 34755 MW; 750C980DF9F94B355 CRC64;

Query Match 63.2%; Score 36; DB 10; Length 300;  
 Best Local Similarity 60.0%; Pred. No. 16;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAMKFFTT 11  
 11:||||  
 DB 40 PAVLNEVTT 49

RESULT 5  
 ID 09XXP5 PRELIMINARY; PRT; 353 AA.  
 AC 09XXP5;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE F08E10.3 PROTEIN.  
 GN F08E10.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gardner A.E.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL: AL021566; CA16499.1; -  
 DR InterPro: IPR003003; 7TM\_chemorecept\_2.  
 DR InterPro: IPR000168; 7TM\_nematode.  
 DR Pfam: PF01604; 7tm\_5; 1.  
 SQ SEQUENCE 353 AA; 39782 MW; 87D126E242BF754E CRC64;

Query Match 63.2%; Score 36; DB 5; Length 353;

Best Local Similarity 60.0%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
OY 2 PAPMKFFTT 11  
11111111  
Db 187 PPTPTEFFTT 196

RESULT 6  
O9AXL7 PRELIMINARY; PRT: 519 AA.  
AC O9AXL7; 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
OS MALATE SYNTHASE (EC 4.1.3.2) (FRAGMENT).  
OC Musa acuminata (Banana).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Zingiberales; Musaceae;  
OC Musa.  
OX NCBI\_TaxID=4641;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV, WILLIAMS;  
RA Chandramouli S., Pua E.C.;  
RT "Molecular cloning and characterization of malate synthase in banana  
(Musa acuminata).";  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: L-MALATE + COA -> ACETYL-COA + H(2)O +  
GLYOXALATE.  
CC -1- PATHWAY: SECOND STEP IN GLYOXALATE BYPASS, AN ALTERNATIVE TO THE  
TRICARBOXYLIC ACID CYCLE (IN BACTERIA, FUNGI AND PLANTS).  
CC -1- SIMILARITY: BELONGS TO THE MALATE SYNTHASE FAMILY.  
DR EMBL; AF321286; AA07428.1; -  
DR InterPro; IPR001465; Malate\_synthase.  
DR PROSITE; PS00510; MALATE\_SYNTHASE.1.  
KM Glyoxylate bypass; Lyase; Tricarboxylic acid cycle.  
FT NON\_TER 1  
SQ SEQUENCE 519 AA; 59224 MW; 854A3CC4B5768661 CRC64;

Query Match 63.2%; Score 36; DB 10; Length 519;  
Best Local Similarity 75.0%; Pred. No. 28;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
OY 4 PMMKFFTT 11  
11111111  
Db 346 PIMKFFTS 353

RESULT 7  
O9SL50 PRELIMINARY; PRT: 121 AA.  
AC O9SL50; 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
GN AT2G17960. PROTEIN.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV, COLUMBIA;  
RA MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
RA Bueli C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
RA Cronin L.A., Shen M., Vankken S.E., Umayam L., Tallon L.J., Gill J.E.,  
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.N., Somerville C.R.,

RA Copenhagen G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
thaliana";  
RL Nature 402:761-768(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV, COLUMBIA;  
RA Lin X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC006201; AAD20130.1; -  
SQ SEQUENCE 121 AA; 14055 MW; F8879F2C984753C0 CRC64;

Query Match 61.4%; Score 35; DB 10; Length 121;  
Best Local Similarity 50.0%; Pred. No. 11;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
OY 2 PAPMKFFTT 11  
11111111  
Db 68 PLPLQOYFTT 77

RESULT 8  
O9H7F5 PRELIMINARY; PRT: 163 AA.  
AC O9H7F5; 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE CDNA: FLJ20974 FIS, CLONE ADS001596.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=ADIPOSE TISSUE;  
RA Kawakami T., Nozuchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,  
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,  
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,  
RA Odayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,  
RA Isogai T., Sugano S.;  
RT "NEO human cDNA sequencing project.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK024627; BAB14937.1; -  
SQ SEQUENCE 163 AA; 17712 MW; 7CD8DB8306DEB1EDC CRC64;

Query Match 61.4%; Score 35; DB 4; Length 163;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 2 PAPMKFFTT 10  
11111111  
Db 41 PAPVGRFTT 49

RESULT 9  
O9CQ18 PRELIMINARY; PRT: 166 AA.  
AC O9CQ18; 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE 1500026D16RIK. PROTEIN.  
GN Mus musculus (Mouse).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE-CEREBELLUM, AND EMBRYO;  
 RX MEDLINE-21085660; PubMed-11217851;  
 RA Kawai J., Shitagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Norioka P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmung L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK018776; BAB31401.1; -;  
 DR EMBL: AK011556; BAB27694.1; -;  
 DR MGD: MGI:1915459; 1500026D16Rik.  
 SQ SEQUENCE 166 AA; 17823 MW; D631CFE252040417 CRC64;

Query Match 61.4%; Score 35; DB 11; Length 166;  
 Best Local Similarity 66.7%; Pred. No. 15;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 PPMKFFTT 10  
 Db 41 PAVDRFFT 49

RESULT 10  
 Q9V243 PRELIMINARY; PRT; 189 AA.  
 AC Q9V243;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE HYPOTHETICAL 20.9 KDA PROTEIN.  
 GN PAB2198.  
 OS Pyrococcus abyssi.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 OX NCBI\_TaxID=29292;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ORSAY;  
 RA Helling R.;  
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome  
 structure and evolution."  
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AJ248283; CAB49155.1; -;  
 DR EMBL: AJ248283; CAB49155.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 189 AA; 20866 MW; 47177FID072EA5B CRC64;

Query Match 61.4%; Score 35; DB 17; Length 189;  
 Best Local Similarity 75.0%; Pred. No. 17;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 PPMKFFTT 11  
 Db 78 PVMKFFAT 85

RESULT 11  
 057989 PRELIMINARY; PRT; 191 AA.

AC 057989;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE HYPOTHETICAL 21.2 KDA PROTEIN PH0251.  
 GN PH0251.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 OX NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OT3;  
 RX MEDLINE-98344137; PubMed-9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3."  
 RL DNA Res. 5:55-76(1998).  
 DR EMBL: AP000001; BAA29323.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 191 AA; 21150 MW; F7CE582B4F6A0BDE CRC64;

Query Match 61.4%; Score 35; DB 17; Length 191;  
 Best Local Similarity 75.0%; Pred. No. 17;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 PPMKFFTT 11  
 Db 78 PVMKFFAT 85

RESULT 12  
 Q08278 PRELIMINARY; PRT; 222 AA.  
 AC Q08278;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CHROMOSOME XV READING FRAME ORF YOL135C.  
 GN MED7 OR YOL135C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Arino J., Casamayor A., Gancedo C., Lafuente M.J.,  
 RA Aldes M., Casas C., Herrero E.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FY1679;  
 RX MEDLINE-97051593; PubMed-8896270;  
 RA Aldes M., Piedrafilta L., Casas C., Casamayor A., Khalid H.,  
 RA Balcells L., Arino J., Herrero E.;  
 RT "Sequence analysis of a 12 801 bp fragment of the left arm of yeast  
 chromosome XV containing a putative 6-phosphofructo-2-kinase gene, a  
 gene for a possible glycopospholipid-anchored surface protein and six  
 other open reading frames."  
 RT Yeast 12:1053-1058(1996).  
 RL EMBL: Z74877; CAA99156.1; -;  
 DR EMBL: X95465; CAA64734.1; -;  
 DR SGD: S0005495; MED7.  
 SQ SEQUENCE 222 AA; 25585 MW; A88301E4EF3223C9 CRC64;



```

Query Match                      61.4%; Score 35; DB 3; Length 222;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PAPMKFFTT 10
   1 1 1 1 1 1
Db 16 PEPYKFFTT 24

RESULT 13
Q9GME2 PRELIMINARY; PRT; 292 AA.
AC Q9GME2:
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-DEC-2001 (TREMblrel. 18, Last sequence update)
DE 17-BETA-HYDROXYSTEROID DEHYDROGENASE TYPE 1 (FRAGMENT).
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.
OX NCBI_TaxID=9483;
RN [1]
RP SEQUENCE FROM N.A.
RA Husein B., Einspanier A.;
RT "Regulation of 17beta-hydroxysteroid dehydrogenases type 1 and type 7
RT during pregnancy in the marmoset monkey.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL: AF272013; AAC01115.2; -.
DR HSSP: P14061; IEDS.
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short.1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; UNKNOW_N.1.
KW Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 292 AA; 31577 MW; A815EE41A51B43FC CRC64;

Query Match                      61.4%; Score 35; DB 6; Length 292;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PAPMKFFTT 11
   1 1 1 1 1 1
Db 226 PKPALRYFTT 235

RESULT 14
Q9SAD4 PRELIMINARY; PRT; 328 AA.
AC Q9SAD4:
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE F3P19.1 PROTEIN.
GN F3P19.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lee J.M., Lenz C.,
RA Liu S., Li J., Kremenetskaia I., Luros J., Ngan I., Gonzalez A.,
RA Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P.,
RA Hansen N., Huizer L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;

```

```

RT "Arabidopsis thaliana chromosome 1 BAC F3P19 sequence.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RA Theologis A.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC007357; AAD31052.1; -.
DR HSSP: O80337; ZGCC.
DR InterPro: IPR001471; AP2-domain.
DR Pfam: PF00847; AP2-domain; 1.
DR PRINTS: PR00367; ETRHSPELENT.
DR ProDom: PD001423; AP2-domain; 1.
DR SMART: SM00380; AP2; 1.
SQ SEQUENCE 328 AA; 36319 MW; 66320534B5FCAE6B CRC64;

Query Match                      61.4%; Score 35; DB 10; Length 328;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PAPMKFFTT 10
   1 1 1 1 1 1
Db 17 PNPTRKFFTT 25

RESULT 15
Q9XX73 PRELIMINARY; PRT; 388 AA.
AC Q9XX73:
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Y102A5C.21 PROTEIN.
GN Y102A5C.21.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gardner A.E.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: AL031627; CAA20962.1; -.
DR InterPro: IPR003003; 7TM_chemorecept_2.
DR InterPro: IPR000168; 7TM_nematode.
DR Pfam: PF01604; 7tm_5; 1.
SQ SEQUENCE 388 AA; 44245 MW; B50BB2CFE2F9B77E CRC64;

Query Match                      61.4%; Score 35; DB 5; Length 388;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PAPMKFFTT 11
   1 1 1 1 1 1
Db 220 PCPTREFFTT 229

```

Search completed: June 11, 2002, 15:54:24  
Job time: 213 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 15:47:49 ; Search time 15.04 Seconds  
(without alignments)  
76.667 Million cell updates/sec

Title: US-09-727-198-5

Perfect score: 57

Sequence: 1 XPAPMMKFFFTX 12

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.71:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	66.7	157	2 T38153	hypothetical prote
2	37	64.9	285	2 AD1992	hypothetical prote
3	36	63.2	225	2 T24052	hypothetical prote
4	36	63.2	300	2 T06691	hypothetical prote
5	36	63.2	353	2 T20580	hypothetical prote
6	36	63.2	485	2 AD2372	hypothetical prote
7	35	61.4	121	2 E84558	hypothetical prote
8	35	61.4	189	2 D75213	hypothetical prote
9	35	61.4	191	2 D71249	hypothetical prote
10	35	61.4	222	2 S66832	hypothetical prote
11	35	61.4	304	2 T16704	hypothetical prote
12	35	61.4	328	1 DEH07	hypothetical prote
13	35	61.4	328	2 G86263	estradiol 17beta-d
14	35	61.4	388	2 T26360	hypothetical prote
15	35	61.4	443	2 H70216	hypothetical prote
16	35	61.4	476	2 H83265	PTS system, cellob
17	35	61.4	810	2 S67050	probable D-alanyl-
18	34	59.6	162	2 D84684	hypothetical prote
19	34	59.6	274	2 T20786	hypothetical prote
20	34	59.6	278	2 A49067	transcription init
21	34	59.6	284	2 B84508	probable Tail-like
22	34	59.6	336	2 T00832	probable Tail-like
23	34	59.6	373	2 E49094	methylinonyl-COA
24	34	59.6	453	2 S18597	tubulin beta chain
25	34	59.6	487	2 E70480	probable sodium/pr
26	34	59.6	931	2 H96527	protein F27J15.16
27	33	57.9	180	2 PS0010	lignin peroxidase
28	33	57.9	180	2 PS0011	lignin peroxidase
29	33	57.9	273	1 B27035	acid phosphatase (

30	33	57.9	325	2 G70334	hypothetical prote
31	33	57.9	371	1 OPJG65	diarypropene pero
32	33	57.9	371	2 JCI168	lignin peroxidase
33	33	57.9	372	1 OPJGAP	lignin peroxidase
34	33	57.9	372	1 OPJGBP	lignin peroxidase
35	33	57.9	372	1 OPJG3P	lignin peroxidase
36	33	57.9	372	2 B32322	lignin peroxidase
37	33	57.9	372	2 A36693	lignin peroxidase
38	33	57.9	372	2 A43638	lignin peroxidase
39	33	57.9	372	2 T04002	lignin peroxidase
40	33	57.9	372	2 A32322	lignin peroxidase
41	33	57.9	372	2 S01028	lignin peroxidase
42	33	57.9	372	2 JH0156	lignin peroxidase
43	33	57.9	372	2 S69246	lignin peroxidase
44	33	57.9	373	2 A27817	lignin peroxidase
45	33	57.9	461	2 A31237	alpha-2C-adrenergic

#### ALIGNMENTS

RESULT 1  
T38153  
hypothetical protein SPAC22A12.13 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T38153  
R:Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
Submitted to the EMBL Data Library, September 1997  
A:Reference number: 221774  
A:Accession: T38153  
A:Status: preliminary; translated from GR/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-157 <PEAN>  
A:Cross-references: EMBL:Z99295; PIDN:CAB16583.1; GSPDB:GN00066; SPDB:SPAC22A12.13  
A:Experimental source: strain 972h-; cosmid c22A12  
C:Genetics:  
A:Gene: SPDB:SPAC22A12.13  
A:Map position: 1

Query Match 66.7%; Score 38; DB 2; Length 157;  
Best Local Similarity 75.0%; Pred. No. 2.2;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 PAPMMKFF 9.  
DB 69 PAPVLKFF 76

RESULT 2  
AD1992  
hypothetical protein all11489 [imported] - Anabaena sp. (strain PCC 7120)  
C:Species: Anabaena sp.  
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
R:Keneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AD1992  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-285 <KOR>  
A:Cross-references: GB:BA000019; PIDN:BA07855.1; PID:q17135310; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all11489

Query Match 64.9%; Score 37; DB 2; Length 285;

Best Local Similarity 60.0%; Pred. No. 6.7;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11  
|||: |||  
Db 60 PAPVORFFRT 69

## RESULT 3

T24052

hypothetical protein R08H2.6 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T24052

R:Matthews, L.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19835

A:Accession: T24052

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-225 <MIL>

A:Cross-References: EMBL:Z81575; PIDN: CAB04632.1; GSPDB: GN00023; CESP: R08H2.6

A:Experimental source: clone R08H2

C:Genetics:

A:Gene: CESP:R08H2.6

A:Map position: 5

Query Match 63.2%; Score 36; DB 2; Length 225;

Best Local Similarity 60.0%; Pred. No. 8.4;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11  
| | : ||||  
Db 77 PCPADFFTT 86

## RESULT 4

T06691

hypothetical protein T17F15.210 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 24-May-2001

C:Accession: T06691

R:Querier, F.; Choisne, N.; Robert, C.; Brotier, P.; Wincker, P.; Cattoico, L.; Artigou

submitted to the Protein Sequence Database, April 1999

A:Reference number: Z15793

A:Accession: T06691

A:Molecule type: DNA

A:Residues: 1-300 <QUB>

A:Cross-References: EMBL:AL049658; GSPDB: GN00061; ATSP: T17F15.210

A:Experimental source: cultivar Columbia; BAC clone T17F15

C:Genetics:

A:Gene: ATSP:T17F15.210

A:Map position: 3

C:Superfamily: *Arabidopsis thaliana* hypothetical protein At2g17920

Query Match 63.2%; Score 36; DB 2; Length 300;

Best Local Similarity 60.0%; Pred. No. 11;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11  
|||: |||  
Db 40 PAPVLFVTT 49

## RESULT 5

T20580

hypothetical protein F08E10.3 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T20580

R:Gardner, A.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z19296

A:Accession: T20580

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-353 <MIL>

A:Cross-References: EMBL:AL021566; PIDN: CAA16499.1; GSPDB: GN00023; CESP: F08E10.3

A:Experimental source: clone F08E10

C:Genetics:

A:Gene: CESP:F08E10.3

A:Map position: 5

A:Introns: 121/1; 332/1

C:Superfamily: *Caenorhabditis* hypothetical protein C4967.2

Query Match 63.2%; Score 36; DB 2; Length 353;

Best Local Similarity 60.0%; Pred. No. 13;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11  
| | : ||||  
Db 187 PCPTORFFTT 196

## RESULT 6

AD2372

hypothetical protein alr4532 [imported] - *Anabaena* sp. (strain PCC 7120)

C:Species: *Anabaena* sp.

A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002

C:Accession: AD2372

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kurita, T.; Sasamoto, S.; Watanabe, A.; Iriugu

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID: 21595285; PMID: 11759840

A:Accession: AD2372

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-485 <KUR>

A:Cross-References: GB:BA000019; PIDN: BAB76231.1; PID: g17133668; GSPDB: GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr4532

Query Match 63.2%; Score 36; DB 2; Length 485;

Best Local Similarity 75.0%; Pred. No. 19;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 APMKFFTT 10  
|||: |||  
Db 207 APLMEFFTT 214

## RESULT 7

E84558

hypothetical protein At2g17960 [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: E84558

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Beilto, M.I.; Town, C.D.; Fujii, C.Y

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: AB4420; MUID: 20083487

A:Accession: E84558

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-121 <STO>

A:Cross-References: GB:AE002093; NID: g44406822; PIDN: AAD20130.1; GSPDB: GN00139

C:Genetics:

A:Gene: At2g17960  
A:Map position: 2

Query Match  
Best Local Similarity 61.4%; Score 35; DB 2; Length 121;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAMMKFFTT 11  
1 : : : : :  
Db 68 PLPLQYFTT 77

RESULT 8  
D75213  
hypothetical protein PAB2198 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999

C:Accession: D75213  
R:Anonymous, Genoscope

Submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: D75213

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-189 <KAM>

A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAR49155.1; PID:e151504

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB2198

Query Match  
Best Local Similarity 61.4%; Score 35; DB 2; Length 189;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAMKFFTT 11  
1 : : : : :  
Db 78 PVMKFFAT 85

RESULT 9  
D71249

hypothetical protein PH0251 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000

C:Accession: D71249

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki

M.; Ohnuki, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137

A:Accession: D71249

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-191 <KAM>

A:Cross-references: GB:AP000001; NID:g3236128; PIDN:BAA29323.1; PID:g3256640

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH0251

Query Match  
Best Local Similarity 61.4%; Score 35; DB 2; Length 191;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAMKFFTT 11  
1 : : : : :  
Db 78 PVMKFFAT 85

RESULT 10  
S66832

hypothetical protein YOL135c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein AOE222; hypothetical protein 00505

C:Species: Saccharomyces cerevisiae

C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 21-Jul-2000

C:Accession: S66832; S72032

R:Arlino, J.; Casamayor, A.; Gamio, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66814

A:Accession: S66832

A:Molecule type: DNA

A:Residues: 1-222 <ARL>

A:Cross-references: EMBL:Z74877; NID:g1420025; PID:e252305; PID:g1420026; MIPS:YOL135

A:Experimental source: strain S288C

R:Aldea, M.; Piedrafito, L.; Casas, C.; Casamayor, A.; Khalid, H.; Balcells, L.; Arin

yeast 12, 1053-1058, 1996

A:Title: Sequence analysis of a 12 801 bp fragment of the left arm of yeast chromosome

protein and six other open reading frames.

A:Reference number: S72030; MUID:97051593

A:Accession: S72032

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-222 <ALD>

A:Cross-references: EMBL:X95465; NID:g1628437; PIDN:CAA64734.1; PID:g1628440

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996

C:Genetics:

A:Map position: 15L

A:Note: YOL135c

Query Match  
Best Local Similarity 66.7%; Score 35; DB 2; Length 222;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAMMKFFTT 10  
1 : : : : :  
Db 16 PPPVKKFFT 24

RESULT 11  
T16704

hypothetical protein R07G3.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T16704

R:Connell, M.

submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid R07G3.

A:Reference number: Z18562

A:Accession: T16704

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-304 <CON>

A:Cross-references: EMBL:U23452; NID:g733564; PID:g733569; PIDN:AAC46751.1; CESP:R07G

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:R07G3.6

A:introns: 26/3

Query Match  
Best Local Similarity 75.0%; Score 35; DB 2; Length 304;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAMKFFTT 11  
1 : : : : :  
Db 235 PLMKMFTT 242

RESULT 12

DEHUE7

estradiol 17beta-dehydrogenase (EC 1.1.1.62) type 1 [validated] - human

N:Alternate names: 17beta-estradiol dehydrogenase; 17beta-hydroxysteroid dehydrogenase  
 C:Species: Homo sapiens (man)  
 C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Dec-2000  
 A:Accession: A36081; S29288; S01654; S43212; A40146; A60440; A21031; A60628; G02067; A37  
 R:Linu-The, V.; Labrie, C.; Smard, J.; Lachance, Y.; Zhao, H.F.; Couet, J.; Leblanc, G.;  
 Mol. Endocrinol. 4, 268-275, 1990  
 A>Title: Structure of two in tandem human 17beta-hydroxysteroid dehydrogenase genes.  
 A:Reference number: A36081; MUID:90231340  
 A:Accession: A36081  
 A:Molecule type: DNA  
 A:Residues: 1-328 <LNU>  
 A:Cross-references: GB:M27138; NID:9181950; PIDN:AAB16941.1; PID:9181951  
 R:Petoketo, H.; Isomaa, V.; Viikho, R.  
 Eur. J. Biochem. 209, 459-466, 1992  
 A>Title: Genomic organization and DNA sequences of human 17beta-hydroxysteroid dehydroge  
 A:Reference number: S29288; MUID:93011163  
 A:Accession: S29288  
 A:Molecule type: DNA  
 A:Residues: 1-328 <PEL>  
 A:Cross-references: EMBL:M84472; NID:9806392; PIDN:AAB16942.1; PID:9177127  
 A:Note: the authors did not translate the codon for residue 1  
 A:Note: the authors translated the codon AAC for residue 233 as Asp  
 R:Petoketo, H.; Isomaa, V.; Maentausta, O.; Viikho, R.  
 FEBS Lett. 239, 73-77, 1988  
 A>Title: Complete amino acid sequence of human placental 17-beta-hydroxysteroid dehydrog  
 A:Reference number: S01654; MUID:89031223  
 A:Accession: S01654  
 A:Molecule type: mRNA  
 A:Residues: 2-328 <PEL2>  
 A:Cross-references: GB:X13440; NID:923364; PIDN:CAA31792.1; PID:923365  
 A:Accession: S43212  
 A:Molecule type: protein  
 A:Residues: 2-24; 53-68; 205-223 <PE2>  
 R:Linu The, V.; Labrie, C.; Zhao, H.F.; Couet, J.; Lachance, Y.; Smard, J.; Leblanc, G.;  
 Mol. Endocrinol. 3, 1301-1309, 1989  
 A>Title: Characterization of cDNAs for human estradiol 17beta-dehydrogenase and assignme  
 A:Reference number: A40146; MUID:89384667  
 A:Accession: A40146  
 A:Molecule type: mRNA  
 A:Residues: 1-328 <HE>  
 A:Cross-references: GB:M27138; NID:9181950; PIDN:AAB16941.1; PID:9181951  
 R:Cast, M.J.; Sims, H.F.; Murdoch, G.L.; Gast, P.M.; Strauss, A.W.  
 Am. J. Obstet. Gynecol. 161, 1726-1731, 1989  
 A>Title: Isolation and sequencing of a complementary deoxyribonucleic acid clone encodin  
 A:Reference number: A60440; MUID:90102621  
 A:Accession: A60440  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 6-68; 'G', 70-91, 'EAWAGAG', 102-158, 160-178, 'P', 181-311, 'A', 313-328 <GAS>  
 R:Murdoch, G.L.; Chin, C.C.; Offord, R.E.; Bradshaw, R.A.; Warren, J.C.  
 J. Biol. Chem. 258, 11460-11464, 1983  
 A>Title: Human placental estradiol 17-beta-dehydrogenase. Identification of a single his  
 A:Reference number: A21031; MUID:84008135  
 A:Accession: A21031  
 A:Molecule type: protein  
 A:Residues: 220-224 <MUR>  
 R:Murdoch, G.L.; Chin, C.C.; Warren, J.C.  
 Biochemistry 25, 641-646, 1986  
 A>Title: Human placental estradiol 17beta-dehydrogenase: sequence of a histidine-bearing  
 A:Reference number: A60628; MUID:86159758  
 A:Accession: A60628  
 A:Molecule type: protein  
 A:Residues: 208-215 <MUR>  
 R:Labrie, F.; Linu-The, V.; Labrie, C.; Berube, D.; Couet, J.; Zhao, H.F.; Gagne, R.; Sim  
 J. Steroid Biochem. 34, 189-197, 1989  
 A>Title: Characterization of two mRNA species encoding human estradiol 17beta-dehydrogen  
 A:Reference number: A37190; MUID:90173172  
 A:Accession: A37190  
 A:Contents: annotation  
 R:Shen, Y.  
 submitted to the EMBL Data Library, August 1995  
 A:Reference number: G09150  
 A:Accession: G02067  
 A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-328 <SHE>  
 A:Cross-references: EMBL:U34879; NID:9975894; PIDN:AAD05019.1; PID:9975895  
 C:Genetics:  
 A:Gene: GDB:HS017B1; EDH17B2; EDHB  
 A:Cross-references: GDB:127970  
 A:Map position: 17q11-17q21  
 A:Introns: 33/1; 89/1; 149/1; 180/2; 239/3  
 C:Function:  
 A:Description: catalyzes the reversible oxidation of estradiol-17beta to estrone usin  
 A:Note: also catalyzes the reversible oxidation of 5-androstenediol to dehydroepiandr  
 C:Superfamily: estradiol 17beta-dehydrogenase; short-chain alcohol dehydrogenase homo  
 C:Keywords: oxidoreductase  
 E:2-328/Product: estradiol 17beta-dehydrogenase #status experimental <MAT>  
 F:4-189/Domain: short-chain alcohol dehydrogenase homology <SADH>  
 F:211,214,222/Active site: His #status experimental

Query Match 61.4%; Score 35; DB 1; Length 328;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 11  
 Db 248 PKPTLRFYTT 257

RESULT 13  
 G86263  
 hypothetical protein P3F9J.1 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 A:Accession: G86263  
 R:Theologias, A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
 ansen, N.F.; Hughes, B.; Hulzar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marzia  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: G86263  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-328 <STO>  
 A:Cross-references: GB:AE005172; NID:94850382; PIDN:AAD31052.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 61.4%; Score 35; DB 2; Length 328;  
 Best Local Similarity 66.7%; Pred. No. 20;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 10  
 Db 17 PNPTRFFTT 25

RESULT 14  
 T26360  
 hypothetical protein Y102A5C.21 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 A:Accession: T26360  
 R:Gardner, A.  
 submitted to the EMBL Data Library, September 1998  
 A:Reference number: Z20204  
 A:Accession: T26360  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-388 <WIL>  
 A:Cross-references: EMBL:AL031627; PIDN:CAA20962.1; CESP:Y102A5C.21  
 A:Experimental source: clone Y102A5C  
 C:Genetics:  
 A:Gene: CESP:Y102A5C.21  
 A:Introns: 154/1; 364/1  
 C:Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 61.4%; Score 35; DB 2; Length 388;  
 Best Local Similarity 60.0%; Pred. NO. 24;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 PAPMAKFFTT 11  
 | | : ||||  
 Db 220 PCPTREFFTT 229

RESULT 15  
 H70216  
 PTS system, cellulose-specific IIC component (celB) homolog - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C>Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 21-Jul-2000  
 C:Accession: H70216  
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
 A:Reference number: A70100; MUID:98065943  
 A:Accession: H70216  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-443 <KLE>  
 A:Cross-references: GB:AE000792; MID:93253098; PIDN:AAC66324.1; PID:92689900; TIGR:BBB04  
 A:Experimental source: strain B31  
 C:Genetics:  
 A:Genome: plasmid  
 C:Superfamily: phosphotransferase system enzyme II factor II, phosphoenolpyruvate-depend

Query Match 61.4%; Score 35; DB 2; Length 443;  
 Best Local Similarity 60.0%; Pred. NO. 28;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 PAPMAKFFTT 11  
 |||: ||:|  
 Db 396 PAPIAFFFT 405

Search completed: June 11, 2002, 15:51:30  
 Job time: 221 sec

**This Page Blank (uspto)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 11, 2002, 15:46:47 ; Search time 13.03 seconds  
(without alignments)  
22.495 Million cell updates/sec

Title: US-09-727-198-5  
Perfect score: 57  
Sequence: 1 XPAPMKKFTTX 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfilltest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	61.4	327	1	US-08-375-962B-12
2	35	61.4	327	2	US-08-562-114B-12
3	35	61.4	327	2	US-08-562-114B-12
4	34	59.6	278	1	US-08-729-594A-12
5	34	59.6	278	1	US-08-188-582-9
6	33	57.9	280	1	US-08-646-715-9
7	33	57.9	311	4	US-08-409-452-1
8	33	57.9	330	5	US-09-056-556-188
9	33	57.9	330	5	PCT-US93-08528-19
10	33	57.9	461	1	US-08-194-338-4
11	32	56.1	229	2	US-08-817-997A-2
12	32	56.1	371	4	US-09-500-569-16
13	32	56.1	1597	4	US-09-423-890-13
14	32	56.1	1597	4	US-08-628-829-14
15	31	54.4	378	4	US-09-500-569-10
16	31	54.4	481	6	5164481-1
17	30	52.6	275	1	US-08-252-995D-13
18	30	52.6	275	2	US-08-834-108-13
19	30	52.6	341	4	US-09-347-803-16
20	30	52.6	382	2	US-08-872-302-2
21	30	52.6	976	3	US-08-560-005-2
22	30	52.6	976	3	US-09-195-868-14
23	30	52.6	976	4	US-09-418-540-2
24	30	52.6	1187	4	US-08-664-962B-8
25	30	52.6	1187	4	US-09-311-743-8
26	30	52.6	1189	3	US-09-195-868-15
27	30	52.6	1229	3	US-09-195-868-28

28	29	50.9	56	1	US-08-358-160-87	Sequence 87, Appl
29	29	50.9	362	4	US-09-500-569-14	Sequence 14, Appl
30	29	50.9	769	3	US-09-320-878-12	Sequence 12, Appl
31	29	50.9	809	4	US-09-105-537-24	Sequence 24, Appl
32	29	50.9	824	2	US-08-785-310A-7	Sequence 7, Appl
33	29	50.9	824	2	US-08-816-693A-52	Sequence 52, Appl
34	29	50.9	824	3	US-08-885-291-52	Sequence 52, Appl
35	29	50.9	824	4	US-09-496-672-52	Sequence 52, Appl
36	29	50.9	1093	3	US-08-545-860D-55	Sequence 55, Appl
37	29	50.9	1093	5	PCT-US94-04496-55	Sequence 55, Appl
38	29	50.9	1185	4	US-09-041-886-23	Sequence 23, Appl
39	29	50.9	3782	4	US-09-105-537-4	Sequence 4, Appl
40	28	49.1	123	4	US-09-500-569-28	Sequence 28, Appl
41	28	49.1	129	3	US-08-942-686-3	Sequence 3, Appl
42	28	49.1	232	1	US-07-797-556-4	Sequence 4, Appl
43	28	49.1	232	1	US-08-225-989-4	Sequence 4, Appl
44	28	49.1	232	1	US-08-570-923-4	Sequence 4, Appl
45	28	49.1	232	1	US-08-580-014-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-08-375-962B-12  
Sequence 12, Application US/08375962B  
Patent No. 5731195  
GENERAL INFORMATION:  
APPLICANT: SIMON, ANDRAS; HELLMAN, ULF; WERNSTEDT,  
APPLICANT: CHRISTER, ERIKSSON, ULF.  
TITLE OF INVENTION: Isolated Nucleic Acid Molecule  
WHICH CODES FOR A 32 kDa Protein Having 11-Cis Retinol  
TITLE OF INVENTION: Dehydrogenase Activity, and Which Associates With p63,  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Felife & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect (ASCII standard)  
CURRENT APPLICATION DATA: US/08375,962B  
APPLICATION NUMBER: US/08375,962B  
FILING DATE: 20-January-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/258,418  
FILING DATE: 6-October-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasqualini, Patricia A.  
REGISTRATION NUMBER: 34,894  
REFERENCE/DOCKET NUMBER: LUD 5372.1 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Human estradiol 17-b dehydrogenase (hEDH)  
US-08-375-962B-12  
Query Match 61.4%; Score 35; DB 1; Length 327;

Best Local Similarity 50.0%; Pred. No. 20;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 11  
| | : : : | | |  
Db 247 PKPTLRFTT 256

RESULT 2  
US-08-562-114B-12

; Sequence 12, Application US/08562114B  
; Patent No. 5972646

; GENERAL INFORMATION:

; APPLICANT: ERIKSSON ET AL.

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH CODES FOR A

; TITLE OF INVENTION: 32 KDA PROTEIN HAVING 11-CIS RETINOL HYDROGENASE

; TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A

; TITLE OF INVENTION: PORTION OF A RETINOL BINDING PROTEIN RECEPTOR

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felte & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect 5.1 and ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/562,114B

; FILING DATE: 22-No. 5972646ember-1995

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/375,962

; FILING DATE: 20-January-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Kohl, Vineet

; REGISTRATION NUMBER: 37,003

; REFERENCE/DOCKET NUMBER: LUD 5372.2 CIP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 327 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FEATURE:

; NAME/KEY: Human estradiol 17-b dehydrogenase (hbdh)

US-08-562-114B-12

Query Match 61.4%; Score 35; DB 2; Length 327;

Best Local Similarity 50.0%; Pred. No. 20;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 11  
| | : : : | | |  
Db 247 PKPTLRFTT 256

RESULT 3

US-08-729-594A-12

; Sequence 12, Application US/08729594A

; Patent No. 6280997

; GENERAL INFORMATION:

; APPLICANT: Eriksson, Ulf; Simon, Andreas; Romert, Anna

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH

; TITLE OF INVENTION: CODES FOR A 32 KDA PROTEIN HAVING 11-CIS RETINOL DEHYDROGENASE

; TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A PORTION OF A

; TITLE OF INVENTION: RETINOL BINDING PROTEIN RECEPTOR

; NUMBER OF SEQUENCES: 41

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felte & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/729,594A

; FILING DATE: 11-October-1996

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/562,114

; FILING DATE: 22-No. 628097ember-1995

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/375,962

; FILING DATE: 20-January-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 628097man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5372.3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 327 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FEATURE:

; NAME/KEY: Human estradiol 17-b dehydrogenase (hbdh)

US-08-729-594A-12

Query Match 61.4%; Score 35; DB 4; Length 327;

Best Local Similarity 50.0%; Pred. No. 20;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 11  
| | : : : | | |  
Db 247 PKPTLRFTT 256

RESULT 4

US-08-188-582-9

; Sequence 9, Application US/08188582

; Patent No. 5534410

; GENERAL INFORMATION:

; APPLICANT: Tjian, Robert

; APPLICANT: Comal, Lucio

; APPLICANT: Dynlacht, Brian D.

; APPLICANT: Hoey, Timothy

; APPLICANT: Ruppert, Siegfried

; APPLICANT: Tanese, Naoko

; APPLICANT: Weinzierl, Robert O.J.

; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,

; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFS AND METHODS OF USE

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

RESULT 5  
 US-08-646-715-9  
 Sequence 9, Application US/08646715  
 Patent No. 5637686  
 GENERAL INFORMATION:  
 APPLICANT: Tjian, Robert  
 APPLICANT: Comai, Lucio  
 APPLICANT: Dynlacht, Brian D.  
 APPLICANT: Hoey, Timothy  
 APPLICANT: Ruppert, Siegfried  
 APPLICANT: Tanese, Naoko  
 APPLICANT: Wang, Edith  
 TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,  
 TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TARS AND METHODS OF USE  
 NUMBER OF SEQUENCES: 36  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: FLEHR, HORBACH, TEST, ALBRITTON & HERBERT  
 STREET: 4 Embarcadero Center, Suite 3400  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-4187  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/646,715  
 FILING DATE: 09-MAY-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/188,582  
 FILING DATE: 28-JAN-1994

RESULT 6  
 US-08-409-452-1  
 Sequence 1, Application US/08409452  
 Patent No. 5747293  
 GENERAL INFORMATION:  
 APPLICANT: DOUGAN, Gordon  
 APPLICANT: FRANKEL, Gad  
 TITLE OF INVENTION: Antibodies to Intimin-like  
 TITLE OF INVENTION: Proteins of E. coli  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: HALE and DORR  
 STREET: 1455 Pennsylvania Avenue, N.W.  
 CITY: Washington  
 STATE: DISTRICT OF COLUMBIA  
 COUNTRY: USA  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/409,452  
 FILING DATE: 23-MAR-1995  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BAKER, Hollie L.  
 REGISTRATION NUMBER: 31,321  
 REFERENCE/DOCKET NUMBER: 102286, 319  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 942-8400  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 280 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-409-452-1

Query Match	57.9%	Score 33;	DB 1;	Length 280;
Best Local Similarity	66.7%	Pred. NO. 42;		
Matches	6;	Conservative	2;	Mismatches
			1;	Indels
			0;	Gaps
0y	3	APMKKEFTT	11	

Db 86 APEVEFTT 94

RESULT 7

US-09-056-556-188

Sequence 188, Application US/09056556

Patent No. 6350456

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, David C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND

NUMBER OF SEQUENCES: 241

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/056,556

FILING DATE: 07-APR-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.457

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 188:

SEQUENCE CHARACTERISTICS:

LENGTH: 311 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-09-056-556-188

Query Match 57.9%; Score 33; DB 4; Length 311;  
Best Local Similarity 50.0%; Pred. No. 47;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFTT 11  
DB 8 PVPMEFLTS 17

RESULT 8

US-08-118-270-19

Sequence 19, Application US/08118270

Patent No. 5508384

GENERAL INFORMATION:

APPLICANT: Murphy, Randall B.

APPLICANT: Schuster, David I.

TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN

TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF

NUMBER OF SEQUENCES: 348

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

TREATM

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/118,270  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236  
FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY-2A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 330 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-118-270-19

Query Match 57.9%; Score 33; DB 1; Length 330;  
Best Local Similarity 62.5%; Pred. No. 50;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFF 9  
DB 283 PGPLFKFF 290

RESULT 9

PCT-US93-08528-19

Sequence 19, Application PC/TUS9308528

GENERAL INFORMATION:

APPLICANT: New York University

TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN

TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF

NUMBER OF SEQUENCES: 348

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/08528

FILING DATE: 09-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/943,236

FILING DATE: 10-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Townsend, Kevin G.

REGISTRATION NUMBER: 34,033

REFERENCE/DOCKET NUMBER: MURPHY-2 PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

TELEX: 248633

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 330 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US93-08528-19

Query Match 57.9%; Score 33; DB 5; Length 330;  
Best Local Similarity 62.5%; Pred. No. 50;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAPMMKFF 9  
1 1 1 1 1  
DB 283 PGPFLKFF 290

RESULT 10  
US-08-194-338-4  
Sequence 4, Application US/08194338  
Patent No. 5474898  
GENERAL INFORMATION:  
APPLICANT: Venter, John C.  
APPLICANT: Fraser, Claire M.  
APPLICANT: McCormick, William R.  
TITLE OF INVENTION: OCTOPAMINE RECEPTOR  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobb, Martens, Olson and Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/194.338  
FILING DATE: 08-FEB-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/676,174  
FILING DATE: 28-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelson, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH101.001DV1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 461 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-08-194-338-4

Query Match 57.9%; Score 33; DB 1; Length 461;  
Best Local Similarity 62.5%; Pred. No. 71;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
OY 2 PAPMMKFF 9  
1 1 1 1 1  
DB 414 PGPFLKFF 421

RESULT 11  
US-08-817-997A-2  
Sequence 2, Application US/08817997A  
Patent No. 5827719  
GENERAL INFORMATION:  
APPLICANT: Sandel, Thomas  
APPLICANT: Kauppinen, Sakari  
APPLICANT: Kofod, Lene V.  
TITLE OF INVENTION: An Enzyme with Lipolytic  
TITLE OF INVENTION: Activity  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 58277190 No. 5827719disk Of No. 5827719th America, Inc.  
STREET: 405 Lexington Avenue - 64th Fl.  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817.997A  
FILING DATE: 24-APR-1997  
CLASSIFICATION: 7435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4316.204  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-878-9652  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-08-817-997A-2

Query Match 56.1%; Score 32; DB 2; Length 229;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 MKFFTT 11  
1 1 1 1 1 1  
DB 1 MKFFTT 6

RESULT 12  
US-09-500-569-16  
Sequence 16, Application US/09500569  
Patent No. 6329204  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Rafalski, Antoni  
APPLICANT: Shen, Jennie  
TITLE OF INVENTION: plant Caffeic acid 3-O-Methyltransferase Homologs  
FILE REFERENCE: BB1327 US NA  
CURRENT APPLICATION NUMBER: US/09/500.569  
CURRENT FILING DATE: 2000-02-09  
EARLIER APPLICATION NUMBER: 60/119,587  
EARLIER FILING DATE: 1999-February-10  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 16

```
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-500-569-16

Query Match
Best Local Similarity 56.1%; Score 32; DB 4; Length 371;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 PAPMKFFT 10
   |||: |||
Db 114 PAPVCRWFT 122

RESULT 13
US-09-423-890-13
; Sequence 13, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES
; FILE REFERENCE: CPT-085CPC
; CURRENT APPLICATION NUMBER: US/09/423,890
; CURRENT FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USSN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USSN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1597
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-423-890-13

Query Match
Best Local Similarity 56.1%; Score 32; DB 4; Length 1597;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PAPMKFFT 11
   ||: ||| |
Db 212 PADRLKFFT 221

RESULT 14
US-08-628-829-14
; Sequence 14, Application US/08628829A
; Patent No. 6333170
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To External
; FILE REFERENCE: CPT-004DVCP3
; CURRENT APPLICATION NUMBER: US/08/628,829A
; CURRENT FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: 08/440,421
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/323,460
; EARLIER FILING DATE: 1994-10-14
; EARLIER APPLICATION NUMBER: 08/049,254
; EARLIER FILING DATE: 1993-05-15
; EARLIER APPLICATION NUMBER: 08/410,602
; EARLIER FILING DATE: 1995-04-24
; EARLIER APPLICATION NUMBER: 08/472,934
; EARLIER FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1597
; TYPE: PRT
```

```
; ORGANISM: Mus musculus
US-08-628-829-14

Query Match
Best Local Similarity 56.1%; Score 32; DB 4; Length 1597;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PAPMKFFT 11
   ||: ||| |
Db 212 PADRLKFFT 221

RESULT 15
US-09-500-569-10
; Sequence 10, Application US/09500569
; Patent No. 6329204
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs
; FILE REFERENCE: BR1327 US NA
; CURRENT APPLICATION NUMBER: US/09/500,569
; CURRENT FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 60/119,587
; EARLIER FILING DATE: 1999-February-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-500-569-10

Query Match
Best Local Similarity 54.4%; Score 31; DB 4; Length 378;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFT 10
   |||: : ||
Db 118 PAPVCRUFT 126
```

Search completed: June 11, 2002, 15:51:20  
Job time: 273 sec



**This Page Blank (uspto)**





PT biodiversity -  
XX  
PS Claim 20; SEQ ID No 39492; 103bp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
SQ Sequence 1332 AA;  
  
Query Match 78.9%; Score 45; DB 22; Length 1332;  
Best Local Similarity 80.0%; Pred. No. 4;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 PAPMKFFTT 11  
| | | : | | | |  
Db 76 papmkfftt 85  
  
RESULT 2  
ABR30037  
ID ABR30037 standard; Peptide; 19 AA.  
XX  
AC ABR30037;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Peptide #2688 encoded by breast cell single exon nucleic acid probe.  
XX  
KW Human; microarray; single exon probe; gene expression; breast;  
KW disease; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157271-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000662.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-496933/54.  
XX

PT New spatially-addressable set of single exon nucleic acid probes,  
PT useful for measuring gene expression in sample derived from human  
PT breast, comprises number of single exon nucleic acid probes -  
PS  
XX Claim 27; SEQ ID NO 13005; 327bp + sequence listing; English.  
XX  
CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BR 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
SQ Sequence 19 AA;  
  
Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 PAPMKFFTT 10  
| | | | | |  
Db 2 purpmsfftt 10  
  
RESULT 3  
ABR35209  
ID ABR35209 standard; Peptide; 19 AA.  
XX  
AC ABR35209;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #2715 encoded by human foetal liver single exon probe.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
KW  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human fetal liver -  
XX Claim 27; SEQ ID NO 27844; 639bp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
XX Sequence 19 AA;  
SQ

Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Db 2 PAPMMKFFT 10  
1 1 1 1 1 1  
2 ptpqmsfft 10

RESULT 4  
ABB20650  
ID ABB20650 standard; Protein: 19 AA.  
XX  
XX ABB20650;  
XX  
XX 23-JAN-2002 (first entry)  
XX  
XX Protein #2649 encoded by probe for measuring heart cell gene expression.  
XX  
XX Human; gene expression: heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX  
XX Homo sapiens.  
XX  
XX WO200157274-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00666.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-488899/53.  
XX  
XX Single exon nucleic acid probes for analyzing gene expression in human  
XX hearts -  
XX  
XX Claim 15; SEQ ID NO 22420; 530pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
XX Sequence 19 AA;  
SQ

Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Db 2 PAPMMKFFT 10  
1 1 1 1 1 1  
2 ptpqmsfft 10

RESULT 5  
AAM56041  
ID AAM56041 standard; Protein: 19 AA.  
XX  
XX AAM56041;  
XX  
XX 05-NOV-2001 (first entry)  
XX  
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 28146.  
XX  
XX Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX  
XX Homo sapiens.  
XX  
XX WO200157275-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00667.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-483446/52.  
XX  
XX Single exon nucleic acid probes for analyzing gene expression in human  
XX brains -  
XX  
XX Example 4; SEQ ID NO: 28146; 650pp + Sequence Listing; English.  
XX  
XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX  
XX Sequence 19 AA;  
SQ

Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFT 10  
| | | | |  
Db 2 ptpqmsfft 10

RESULT 6  
AAM68412  
ID AAM68412 standard; Protein; 19 AA.  
XX  
AC AAM68412;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 28718.  
XX  
KM Human: bone marrow expressed exon; gene expression analysis; probe;  
XX microarray; cancer; leukaemia; lymphoma; myeloma.  
XX  
OS Homo sapiens.  
XX  
PN WO200157276-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00668.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488901/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human bone marrow -  
XX  
XX  
PS Example 4; SEQ ID NO: 28718; 658bp + Sequence Listing; English.  
XX  
XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.  
XX  
SQ Sequence 19 AA;

Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFT 10  
| | | | |  
Db 2 ptpqmsfft 10

RESULT 7  
AAM16227  
ID AAM16227 standard; Protein; 19 AA.  
XX  
AC AAM16227;  
XX  
DT 12-OCT-2001 (first entry)

XX  
DE Peptide #2661 encoded by probe for measuring cervical gene expression.  
XX  
KM Probe: human; microarray; gene expression; cervical epithelial cell;  
XX cervical cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157278-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00670.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488901/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human cervical epithelial cells -  
XX  
XX  
PS Claim 27; SEQ ID No 21053; 487bp; English.  
XX  
XX The present invention relates to human single exon nucleic acid probes  
CC (SENPs: see AAI10068-AA128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 19 AA;

Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFT 10  
| | | | |  
Db 2 ptpqmsfft 10

RESULT 8  
AAM28717  
ID AAM28717 standard; Protein; 19 AA.  
XX  
AC AAM28717;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Peptide #2754 encoded by probe for measuring placental gene expression.  
XX  
KM Probe: microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.

XX 09-AUG-2001.  
PD Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI: 2001-488897/53.  
PF 30-JAN-2001; 2001WO-US00663.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI: 2001-488897/53.  
DR  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
PS Claim 27; SEQ ID No 28986; 654bp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENPs;  
CC see AA13115-AA157346). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.  
XX  
SQ Sequence 19 AA:  
  
Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 PAPMMKFFT 10  
DB 2 ptpqmsfft 10  
| | | | |  
| | | | |  
  
RESULT 9  
AA03957  
ID AA03957 standard; Protein; 19 AA.  
XX  
XX AA03957;  
AC  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE Peptide #2639 encoded by probe for measuring breast gene expression.  
XX  
KW Probe; human; breast disease; breast cancer; development disorder;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
XX WO200157270-A2.  
PN  
XX  
PD 09-AUG-2001.  
XX  
PF 29-JAN-2001; 2001WO-US00661.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI: 2001-476286/51.  
DR  
XX  
PT Novel single exon nucleic acid probe used to measuring gene expression  
PT in a human breast -  
XX  
PS Claim 27; SEQ ID No 12697; 322pp; English.  
XX  
CC The present invention relates to novel single exon nucleic acid probes  
CC (see AA100010-AA110067). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for measuring human gene expression in  
CC a human breast sample, where the probe hybridises at high stringency to a  
CC nucleic acid expressed in the human breast. The probes are useful for  
CC predicting, diagnosing, grading, staging, monitoring and prognosis  
CC diseases of the human breast, particularly those diseases with polygenic  
CC aetiology. The diseases include: breast cancer, disorders of development,  
CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
CC breast disease and non-carcinoma tumours.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 19 AA:  
  
Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 PAPMMKFFT 10  
DB 2 ptpqmsfft 10  
| | | | |  
| | | | |  
  
RESULT 10  
AA012771  
ID AA012771 standard; Protein; 140 AA.  
XX  
XX AA012771;  
AC  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SPQ ID NO: 12012.  
XX  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132486.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144614.

PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158322.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Query Match

Best Local Similarity 63.2%; Score 36; DB 21; Length 140;  
Pred. No. 23;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 APMKKFFTT 11  
: | : |||||  
Db 27 spykfftt 35

## RESULT 11

ABB63911  
ID ABB63911 standard; Protein: 2016 AA.

XX ABB63911;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 18525.

KW Drosophila: developmental biology; cell signalling; insecticide;

KM pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

DR N-PSDB; ABL08014.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -

XX Disclosure; SEQ ID NO 18525; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins

CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 2016 AA;

Query Match 61.4%; Score 35; DB 22; Length 2016;  
Best Local Similarity 50.0%; Pred. No. 4.7e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 PPMKKFFTT 11  
: | : |||||  
Db 1118 pspqlsft 1127

## RESULT 12

AA002149  
ID AA002149 standard; Protein: 145 AA.

XX AA002149;

XX 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 16041.

KW Human: cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haemotopiasis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RF;

XX WPI: 2001-514838/56.

DR N-PSDB; AAI82080.

PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -

XX Claim 20; SEQ ID NO 16041; 1399pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AA179941-AA193841) and  
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 145 AA;

Query Match 59.6%; Score 34; DB 22; Length 145;  
Best Local Similarity 66.7%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 PAPMMKFFT 10  
I I I I I I  
Db 25 PPPMkffs 33  
RESULT 13  
AAG38411  
ID AAG38411 standard; Protein: 155 AA.  
XX  
AC AAG38411;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47383.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PF 25-FEB-1999; 990S-0121825.  
XX PR 05-MAR-1999; 990S-0123180.  
PR 09-MAR-1999; 990S-0125348.  
PR 23-MAR-1999; 990S-0125788.  
PR 25-MAR-1999; 990S-0126264.  
PR 29-MAR-1999; 990S-0126785.  
PR 01-APR-1999; 990S-0127462.  
PR 06-APR-1999; 990S-0128234.  
PR 08-APR-1999; 990S-0128714.  
PR 16-APR-1999; 990S-0129845.  
PR 19-APR-1999; 990S-0130077.  
PR 21-APR-1999; 990S-0130449.  
PR 23-APR-1999; 990S-0130510.  
PR 23-APR-1999; 990S-0130891.  
PR 28-APR-1999; 990S-0131449.  
PR 30-APR-1999; 990S-0132048.  
PR 30-APR-1999; 990S-0132407.  
PR 04-MAY-1999; 990S-0132484.  
PR 05-MAY-1999; 990S-0132485.  
PR 06-MAY-1999; 990S-0132486.  
PR 06-MAY-1999; 990S-0132487.  
PR 07-MAY-1999; 990S-0132863.  
PR 11-MAY-1999; 990S-0134256.  
PR 14-MAY-1999; 990S-0134218.  
PR 14-MAY-1999; 990S-0134219.  
PR 14-MAY-1999; 990S-0134221.  
PR 14-MAY-1999; 990S-0134370.  
PR 18-MAY-1999; 990S-0134768.  
PR 19-MAY-1999; 990S-0134941.  
PR 20-MAY-1999; 990S-0135124.  
PR 21-MAY-1999; 990S-0135353.  
PR 24-MAY-1999; 990S-0135629.  
PR 25-MAY-1999; 990S-0136021.  
PR 27-MAY-1999; 990S-0136392.  
PR 28-MAY-1999; 990S-0136782.  
PR 01-JUN-1999; 990S-0137222.  
PR 03-JUN-1999; 990S-0137528.  
PR 04-JUN-1999; 990S-0137502.  
PR 07-JUN-1999; 990S-0137794.  
PR 08-JUN-1999; 990S-0138094.  
PR 10-JUN-1999; 990S-0138540.

PR 10-JUN-1999; 990S-0138847.  
PR 14-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139452.  
PR 16-JUN-1999; 990S-0139453.  
PR 17-JUN-1999; 990S-0139492.  
PR 18-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139455.  
PR 18-JUN-1999; 990S-0139456.  
PR 18-JUN-1999; 990S-0139457.  
PR 18-JUN-1999; 990S-0139458.  
PR 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 18-JUN-1999; 990S-0139763.  
PR 21-JUN-1999; 990S-0139817.  
PR 22-JUN-1999; 990S-0139899.  
PR 23-JUN-1999; 990S-0140353.  
PR 23-JUN-1999; 990S-0140354.  
PR 24-JUN-1999; 990S-0140695.  
PR 28-JUN-1999; 990S-0140823.  
PR 29-JUN-1999; 990S-0140991.  
PR 30-JUN-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141842.  
PR 01-JUL-1999; 990S-0142154.  
PR 02-JUL-1999; 990S-0142055.  
PR 06-JUL-1999; 990S-0142390.  
PR 08-JUL-1999; 990S-0142803.  
PR 09-JUL-1999; 990S-0142920.  
PR 12-JUL-1999; 990S-0142977.  
PR 13-JUL-1999; 990S-0143542.  
PR 14-JUL-1999; 990S-0143624.  
PR 15-JUL-1999; 990S-0144005.  
PR 16-JUL-1999; 990S-0144085.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 19-JUL-1999; 990S-0144335.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144632.  
PR 20-JUL-1999; 990S-0144884.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 21-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145085.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145089.  
PR 22-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 23-JUL-1999; 990S-0145224.  
PR 26-JUL-1999; 990S-0145276.  
PR 27-JUL-1999; 990S-0145913.  
PR 27-JUL-1999; 990S-0145918.  
PR 27-JUL-1999; 990S-0145919.  
PR 28-JUL-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.  
PR 02-AUG-1999; 990S-0146388.  
PR 02-AUG-1999; 990S-0146389.  
PR 03-AUG-1999; 990S-0147038.  
PR 04-AUG-1999; 990S-0147204.  
PR 04-AUG-1999; 990S-0147302.  
PR 05-AUG-1999; 990S-0147192.  
PR 05-AUG-1999; 990S-0147260.  
PR 06-AUG-1999; 990S-0147303.  
PR 06-AUG-1999; 990S-0147416.  
PR 09-AUG-1999; 990S-0147493.  
PR 09-AUG-1999; 990S-0147935.



PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158222.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 59.68; Score 34; DB 21; Length 155;  
Best Local Similarity 60.08; Pred. No. 60;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 2 PAPMMKFFTT 11  
| | :||||

Db 9 prpshfflt 18  
RESULT 14  
AAG23777  
ID AAG23777 standard; Protein: 159 AA.  
XX  
XX AAG23777;  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27209.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
OS Arabidopsis thaliana.  
PN  
XX  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0142977.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147202.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 59 6%; Score 34; DB 21; Length 159;  
Best Local Similarity 60.0%; Pred. No. 62;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PADMKEFTT 11  
| | :|||  
Db 9 prpshriftt 18

RESULT 15  
AAG32031  
ID AAG32031 standard; Protein; 162 AA.  
XX



PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161922.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match		59.6%;	Score 34;	DB 21;	Length 162;
Best Local Similarity		60.0%;	Pred. No. 63;		
Matches	6; Conservative	1;	Mismatches	3;	Indels 0; Gaps 0;
QY	2 PAPMKFFTT 11				
	1 1 :1111				
DB	9 prpshfflt 18				

Search completed: June 11, 2002, 15:50:48  
Job time: 867 sec